

# Embedded FIR filter Design for Real-Time Refocusing Using a Standard Plenoptic Video Camera

Christopher Hahne and Amar Aggoun

Dept. of Computer Science, University of Bedfordshire,  
Park Square, Luton, Bedfordshire, LU1 3JU, United Kingdom

## ABSTRACT

A novel and low-cost embedded hardware architecture for real-time refocusing based on a standard plenoptic camera is presented in this study. The proposed layout design synthesizes refocusing slices directly from micro images by omitting the process for the commonly used sub-aperture extraction. Therefore, intellectual property cores, containing switch controlled Finite Impulse Response (FIR) filters, are developed and applied to the Field Programmable Gate Array (FPGA) XC6SLX45 from Xilinx. Enabling the hardware design to work economically, the FIR filters are composed of stored product as well as upsampling and interpolation techniques in order to achieve an ideal relation between image resolution, delay time, power consumption and the demand of logic gates. The video output is transmitted via High-Definition Multimedia Interface (HDMI) with a resolution of 720p at a frame rate of 60 fps conforming to the HD ready standard. Examples of the synthesized refocusing slices are presented.

**Keywords:** Light field, plenoptic camera, refocusing, ray tracing, hardware, signal processing, FPGA, FIR filter

## 1. INTRODUCTION

Recent developments in the subject area of light field have led to a renewed interest in digital refocusing which allows the focus of an image to be changed after a capture has been taken. So far, the refocusing method using a standard plenoptic camera has only been applied to single photos. The contribution of the present research goes beyond that, aiming to provide digitally refocused motion pictures to low-cost computing devices. However, the developed hardware design can certainly be utilized for refocusing still plenoptic pictures in real-time as well. Contrary to a conventional camera, the plenoptic camera captures the information of images from different viewpoints by using only one image sensor that has an array of micro lenses in front of it. The obtained data enables synthesis of two-dimensional images with a variable focus after the photo has been taken. More precisely, the objective of this technology is to increase the depth of field by digitally refocusing on objects at different distances from the camera.

The initial concept of a plenoptic camera, consisting of an array of pinholes, was introduced by Ives<sup>1</sup> in 1903. Subsequent research was conducted by Lippmann<sup>2</sup> whose alternative imaging system was the first containing a micro lens array (MLA). Since then, a considerable amount of literature has been published on integral and light field photography.<sup>3-7</sup> In 2000, Isaksen<sup>8</sup> elaborated a digital synthesis technique enabling variation of the focus within a light field, acquired by using an array of cameras. An additional contribution has been carried out by Ng<sup>9</sup> in 2005 when he examined the digital signal processing of data obtained by a standard plenoptic camera. The very first work concerned with a hardware-based design able to process light field data appeared in 2012 and was proposed by Wimalagunaratne.<sup>10</sup> In this publication a systolic array architecture is described employing Infinite Impulse Response (IIR) filters to render light field slices from a 4 by 4 camera array. Although extensive research has been undertaken on that, no single study exists which adequately covers a hardware implementation to synthesize refocusing slices from image content captured by a plenoptic camera. Therefore, the objective of this paper seeks to address that gap and proposes an architecture for a two-dimensional slice synthesis of light field data gathered by a standard plenoptic camera.

---

Further author information: (Send correspondence to C.Hahne or A.Aggoun)

C.Hahne: E-mail: christopher.hahne[at]study.beds.ac.uk, Telephone: +44 (0) 770 646 1446

A.Aggoun: E-mail: amar.aggoun[at]beds.ac.uk, Telephone: +44 (0) 1582 489230

## 2. THEORETICAL CONCEPT OF THE REFOCUSING SYNTHESIS

As pioneered by Levoy,<sup>7</sup> a ray in the light field  $L$  can be parameterized by intersecting two planes in space consisting of two dimensions each, thus four in total. Subsequently advanced by Ng,<sup>9</sup> in the case of the plenoptic camera, the irradiance  $I_B$  at the micro lens array is given by

$$I_B(s, t) = \frac{1}{B^2} \iint L_B(s, t, U, V) A(U, V) \cos^4 \theta \, dU \, dV, \quad (1)$$

where  $(s, t)$  represents all micro lens centers in a two-dimensional array,  $(U, V)$  is the plane of the main lens and  $B$  denotes the distance between image sensor and the exit pupil of the main lens. By assuming the aperture  $A(U, V)$  of the main lens to be fully open, the scale factor is set to  $A(U, V) = 1$  and can therefore be neglected. In addition, vignetting will not be taken into account in the upcoming models, so that the fall-off factor  $\cos^4$  and its incidence angle  $\theta$  can be left out as well. A further assumption to simplify the function is to ignore the constant  $1/B^2$  and to break it down into horizontal dimension, so that follows

$$I_B(s) = \int L_B(s, U) \, dU. \quad (2)$$

In the equation provided by Ng, the term  $I_B$  represents the incident irradiance impinging on the plane of the MLA. However, in fact, the irradiance of the light field is measured by a sensor at the image plane of the micro lenses. Since light is traveling through a micro lens  $s$ , its irradiance  $I_B(s)$  is distributed over the related image plane, denoted as  $u$ , as shown in Figure 1. The image plane domain of one micro lens is a micro image  $W_s(u)$  being a function of  $u$ . Furthermore, for a particular point in  $s$ , the irradiance occurring along the main lens plane  $U$  is weightily projected onto the sensor in one micro image which can be mathematically proven by using the method of similar triangles.

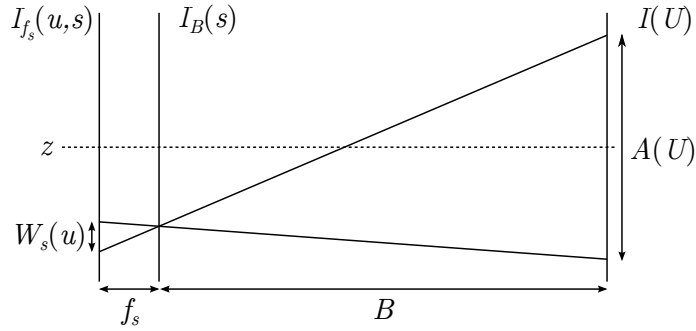


Figure 1. Planes of irradiance

Therefore, the direction of light rays is now determined by its intersection with the sensor and the respective micro lens. Hence, it turns out that the overall irradiance  $I_{f_s}$  at the image plane behind the entire MLA is a function depending on  $(u, s)$ . By disregarding the apertures  $A(s, t)$  of the micro lenses, their vignetting factor  $\cos^4 \theta$  and the constant of the separation between sensor and MLA, the relationship between  $I_B$  and  $I_{f_s}$  is defined as

$$I_B(s) = \int I_{f_s}(u, s) \, du. \quad (3)$$

As the captured wavelength spectrum of the irradiance is assumed to be bandlimited and weighted according to the luminosity function, thus the irradiance can be replaced by the illuminance  $E$  so that

$$E_B(s) = \int E_{f_s}(u, s) \, du. \quad (4)$$

Several attempts have been made to investigate light rays traveling through plenoptic camera setups.<sup>11–13</sup> These studies explain the light field theory mainly using the method of similar triangles, either by simplifying the model to the range from the main lens to the sensor or by beginning from the real object space. Alternatively, it might be an applicable approach to trace rays from the inside of the camera into the object space. Starting from the sensor, the last physical barrier that light rays pass through is the MLA. According to the optical setup of the standard plenoptic camera, the micro lenses are located one focal length from the sensor.<sup>12</sup> By considering the thin lens equation (5), it is shown in (6) and (7) that light rays converging at the distance  $b_s$  being one focal length  $f_s$  behind the micro lens  $s$ , have to be emitted from an approximate infinite distance  $a_s$ . Looking at it from the other perspective, light rays of an emitting point at  $b_s$  from  $s$  will propagate in that manner that they coincide at an image point at infinite distance. As a consequence, rays of a light beam coming from infinity can be seen as traveling parallel, respectively collimated, as they never intersect each other.

$$\frac{1}{f_s} = \frac{1}{a_s} + \frac{1}{b_s} \quad (5)$$

$$\frac{1}{f_s} = \frac{1}{a_s} + \frac{1}{f_s} \quad (6)$$

$$0 = \lim_{a_s \rightarrow \infty} \left( \frac{1}{a_s} \right) \quad (7)$$

$$a_s \rightarrow \infty \quad (8)$$

For this study it is believed that points  $u_i$  along the image plane have some spacing and are hypothetically of an infinitesimal width in the range of  $i = [0, m - 1]$  where  $i \in \mathbb{Z}$  and  $m$  is the total number of samples  $i$  along the  $u$  domain. Similarly, the micro lens plane  $s$  is assumed to be discretized by separated micro lenses  $s_j$  where  $j \in \mathbb{Z}$  numbering the micro lenses consecutively. On this supposition, it is depicted in Figure 2 that these points were formed from light bundles of a corresponding angle consisting of parallel light rays. For clarity, Figure 2 shows only chief rays of light beams passing through the micro lens center.

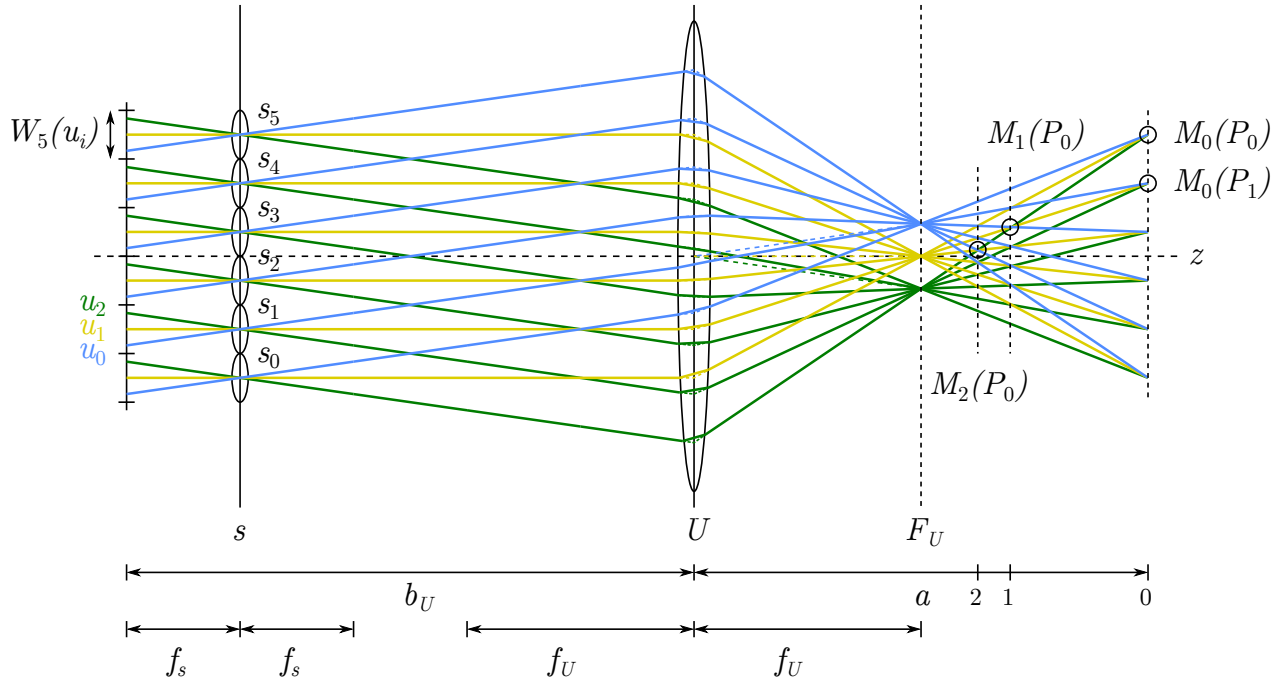


Figure 2. Ray tracing intersection model

Assuming that there is a surface of a Lambertian object in Figure 2 within the range of  $P_0$  at plane  $a = 0$ , light rays are reflected from that position with the luminous emittance  $M_0(P_0)$  in all possible directions, including into the camera, and distributed along  $u_i$  under several micro lenses  $s_j$ . In order to recover the imaginary illuminance  $E'_0(P_0)$ , believed to be equal to  $M_0(P_0)$ , which would have been measured with a conventional camera, ideally, the weighted average of the integrated illuminance  $E_{f_s}[u_i, s_j]$  has to be calculated in the way that

$$E'_0(P_0) = \frac{1}{3} \left( E_{f_s}[u_2, s_0] + E_{f_s}[u_1, s_0] + E_{f_s}[u_0, s_0] \right). \quad (9)$$

Similarly, a direct neighbor point  $P_1$  of the same plane  $a = 0$  is processed in

$$E'_0(P_1) = \frac{1}{3} \left( E_{f_s}[u_2, s_1] + E_{f_s}[u_1, s_1] + E_{f_s}[u_0, s_1] \right), \quad (10)$$

and the illuminance of the point  $M_1(P_0)$  at plane  $a = 1$  being closer to the camera is as follows

$$E'_1(P_0) = \frac{1}{3} \left( E_{f_s}[u_2, s_0] + E_{f_s}[u_1, s_1] + E_{f_s}[u_0, s_2] \right). \quad (11)$$

In order to focus even closer to the camera, an intersection  $P_0$  at plane  $a = 2$  is given by

$$E'_2(P_0) = \frac{1}{3} \left( E_{f_s}[u_2, s_0] + E_{f_s}[u_1, s_2] + E_{f_s}[u_0, s_4] \right). \quad (12)$$

Thus, weighting and summing of appropriate points measured at the sensor allows for the recreation of the illuminance of a plane within the light field which is regarded as refocusing. The number of pixels involved to form a new pixel  $E'_a[P_j]$  equals the micro image resolution  $m$  being therefore the divider for the averaging. Taking into account equation (4), a more general formula can be derived from the given examples to satisfy all intersections of any plane  $a$  as

$$E'_a[P_j] = \frac{1}{m} \sum_{i=0}^{m-1} E_{f_s}[u_{m-1-i}, s_{j+i \times a}], \quad a \in \mathbb{Z}. \quad (13)$$

By applying this equation, it is possible to synthesize two-dimensional images within a light field captured by a standard plenoptic camera system. However, it is necessary to mention that in case of absence of an object surface at the corresponding point of the particular plane in object range, artifacts do occur. This is due to the nature of the sampled light field as the selected points  $u_i$  might carry information of rays emitted from varying depth, respectively objects at different distances. Consequently, the selection and integration of illuminated points  $E_{f_s}[u, s]$  representing rays that never started from the same position  $P$  results in corrupted pixel values in  $E'$ . In other words, the origin of rays in object space varies among points  $u_i$ , hence some possible combinations for  $E'_a$  might lead to high frequency artifacts instead of a blur as in traditional cameras. To overcome this issue, a depth map acquired prior to the refocusing process can be used as an aid to process only those points of a plane where light is reflected from an object. Publications regarding the depth map calculation may be considered for further information.<sup>14,15</sup> Artifacts are not covered in this work and, therefore, are accepted, as the real-time refocusing process takes top priority.

Since the  $E_{f_s}[u_i, s_j]$  domain is defined to be discrete, the light field can be translated into a horizontal plenoptic image  $E_{f_s}[x]$  as it appears on the sensor (see Figure 3) containing pixels  $x$  as shown in

$$x = i + j \times m. \quad (14)$$

Equally, the vertical one-dimensional coordinate  $y$  of a plenoptic image is obtained by

$$y = c + d \times m, \quad (15)$$

where  $c \in \mathbb{Z}$  and  $d \in \mathbb{Z}$  are indices of  $v$  and  $t$  respectively.

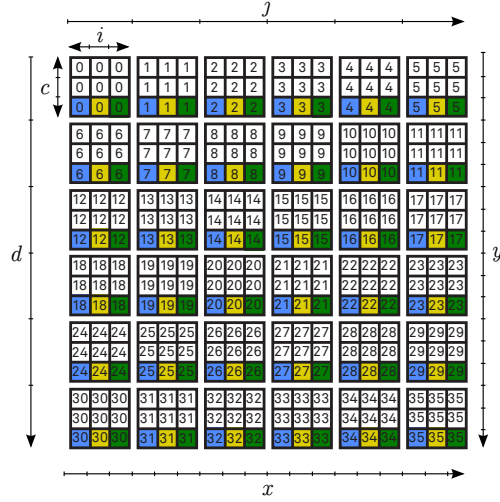


Figure 3. Micro image acquisition model

Note that colored pixels in Figure 3 correspond to the colors representing rays at particular positions  $u_i$  in the ray tracing intersection scheme in Figure 2. In using the formula to translate  $[u, s]$  to the  $x$ -dimension, the equations computing points  $M_a(P)$  given above can be rewritten as

$$E'_0(P_0) = \frac{1}{3} \left( E_{f_s}[2] + E_{f_s}[1] + E_{f_s}[0] \right), \quad (16)$$

$$E'_0(P_1) = \frac{1}{3} \left( E_{f_s}[5] + E_{f_s}[4] + E_{f_s}[3] \right), \quad (17)$$

$$E'_1(P_0) = \frac{1}{3} \left( E_{f_s}[2] + E_{f_s}[4] + E_{f_s}[6] \right), \quad (18)$$

$$E'_2(P_0) = \frac{1}{3} \left( E_{f_s}[2] + E_{f_s}[7] + E_{f_s}[12] \right). \quad (19)$$

In using the standard plenoptic setup and its corresponding image processing, it has been stated that the spatial resolution of a synthesized refocused image equals the number of micro lenses.<sup>13</sup> However, the drawback of low spatial resolution in the standard plenoptic camera may be compensated by taking advantage of upsampling and interpolation. Considering plane  $a = 0$  in Figure 2, it is assumed the resulting refocusing slice has at least  $m$ -times less resolution in respect of the native sensor resolution. As in synthesizing plane  $a = 0$ , the pixels of each micro image are merged, the resolution of a refocused image equals the number of micro images. Therefore, an upsampling factor has to be equal to  $m$  in order to compensate for the resolution loss and approaching the original size of the incoming frame. Since the synthesized resolution is determined by the number of micro images, the upsampling process can be viewed as an interpolation of micro images. Hence, it is now possible to shift by interpolated pixels, called sub-pixels, so that  $a = a'/m$  where  $a'$  denotes the sub-pixel shift.

### 3. HARDWARE DESIGN

In order to process continuous frames within the camera device in real-time, it requires an embedded computing system. Therefore, an FPGA is utilized since it is reconfigurable, low-power-consuming and able to process data in parallel. For the optical calibration of our camera, an autocollimator has been used to align the distance between MLA and image sensor being one focal length behind the MLA.<sup>16</sup> As the array of micro lenses is arranged rectangularly, the main lens has a square aperture attached in front of it in order to obtain square micro images improving the fill factor.<sup>17</sup>

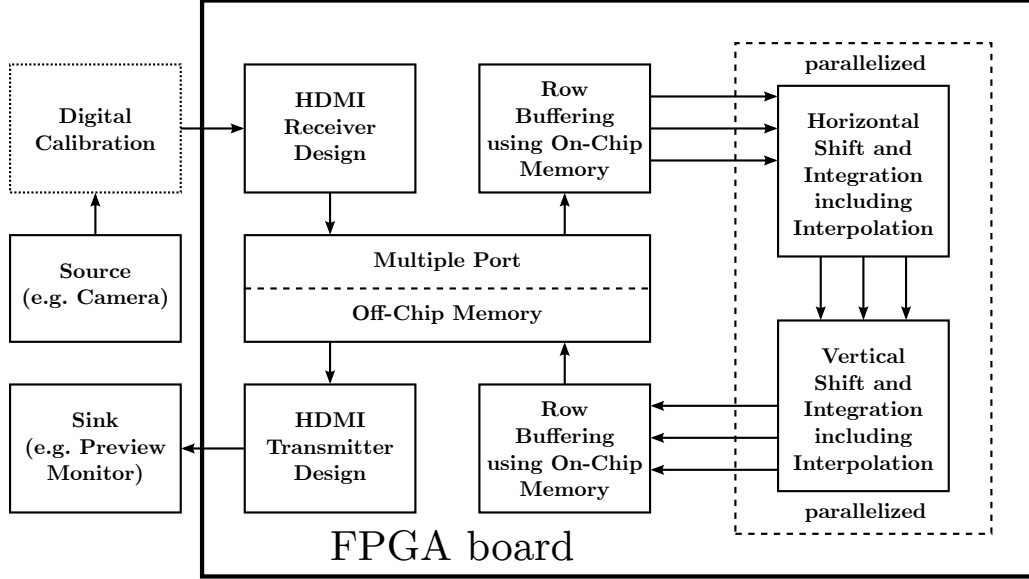


Figure 4. Block diagram for signal processing

Figure 4 outlines the signal processing chain which will be discussed in the following. A digital micro image calibration covers the centroid calculation, rotation and cropping of each micro image. As the digital calibration is considered to be preprocessed, it is not within the scope of this paper to propose hardware implementation attempts for this purpose. The concept for a calibration routine on an FPGA is left as future work. The interested reader may be referred to corresponding literature of that particular field.<sup>16,18–20</sup>

For the purpose of streaming real-time video data, the High-Definition Multimedia Interface (HDMI) was chosen which can be easily applied to a large amount of mobile devices using the Mobile High-Definition Link (MHL) which is a derivate of the HDMI standard.<sup>21</sup> According to the HDMI specification,<sup>22</sup> pixels are transmitted in a serial manner from the source to the sink. Due to the serial transmission of each line, the shift and integration in the horizontal direction might be done before the incoming data is stored in the memory for the first time. However, the horizontal shift and integration can be easily parallelized when it is already stored with the aid of row buffering<sup>23</sup> by synthesizing each line independently and most important simultaneously. Moreover, in this section, it is shown that the parallelization of horizontal processors enables implementation of the vertical shift and integration without necessity of additional off-chip memory. For that reason, the shift and integration is subdivided into a horizontal and vertical process in the proposed hardware design. Hereafter, the main focus embraces these two processing stages. For convenience, the bit depth and colour channels of pixels are disregarded. Thus, in the upcoming drawings each pixel is covered by an FIR filter unit, e.g. a delay register  $z^{-1}$ .

### 3.1 Shift and integration

Assuming the light field  $L(s, t, u, v)$  to be a linear time-invariant system, the shift and integration can be described as a convolution. As shown in equation (20), a discrete FIR filter basically consists of multiplication and addition operations where  $p[n]$  represents the incoming,  $o[n]$  the outgoing digital data samples and  $h[k]$  denotes the dataset of the impulse response. As deduced from the previous section, the illuminance values of several pixels, following the scheme shown in Figure 2, have to be averaged which requires the division operation. Evidently, the division can be seen as a multiplication with a fraction. Hence, these fractions are the filter coefficients  $h[k]$  in

$$o[n] = \sum_{k=0}^{N-1} h[k] \times p[n-k]. \quad (20)$$

Since the fraction for the incoming pixel data  $p[n]$  depends on the given horizontal micro image resolution  $m$  being the same for all micro images after cropping, the impulse response  $h$  for each incoming pixel is  $1/m$  and thus the constant  $h_0$ . By plugging in the denotations defined earlier, it follows

$$E'_{(a'/m)}[x] = \sum_{k=0}^{m-1} \frac{1}{m} \times E_{fs}[x-k]. \quad (21)$$

The so-called boxcar filter, having a constant filter coefficient  $h_0 = 1/m$ , simplifies the multiplication tremendously, as the hardware design can benefit from the sorted product technique. This technique has advantages over the more complex multiplication approach in combinational logic gates and power consuming digital signal processors (DSP's). When using the sorted product method, the results of the multiplication are generated in advance and stored in Read-Only-Memories (ROM's) as Look-Up-Tables (LUT's). The products can be read out by addressing them with the corresponding pixel value  $x$ . In consequence, this process requires only one pixel clock cycle to make the product available which means less delay time. Additionally, this approach saves many logic gates and therefore power consumption compared to the common multiplication in logic gates.

#### 3.1.1 Horizontal shift and integration

According to the ray tracing intersection model, a specific example for the horizontal processing is the differential equation of a 2nd-order FIR filter given by

$$E'_{1/3}[x] = \frac{1}{3} \left( E_{fs}[x-0] + E_{fs}[x-1] + E_{fs}[x-2] \right), \quad (22)$$

where  $N = m = 3$ . For the filter arrangement a systolic design is considered. As systolic arrays inherently broadcast input values  $x$  to several tap registers, it might be applicable to exploit a systolic FIR filter arrangement for nearest-neighbor interpolation. The nearest-neighbor interpolation has been chosen for the sake of simplicity as it basically repeats an incoming pixel value. In Figure 5 a semi-systolic FIR filter for sub-pixel shift  $a = 1/3$  is shown. Note that the depicted FIR design refers to the example stated in the ray tracing model, but the shift of one sub-pixel is not equivalent to a shift of one pixel. The data flow diagram in Figure 6 visualizes the difference by illustrating the combinations of pixels obtained by the camera sensor.

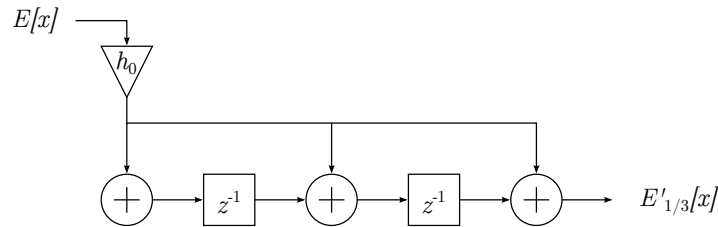


Figure 5. Semi-systolic FIR filter diagram for horizontal sub-pixel shift  $a = 1/3$

Left-sidedly depicted in Figure 6, you find data of one-dimensional serial micro images  $W_{sj}(u_i)$  coming from the video source. Note that the fraction process is not considered in this drawing. The intermediate stage in the middle simply repeats each pixel in a micro image by broadcasting it  $m$  times. Subsequently, the data sets of these interpolated values are summed up at each clock cycle to form the output pixel  $E'_{1/3}[x]$ . Accordingly, this proposition makes an additional process of the frequently used sub-aperture extraction<sup>24, 25</sup> redundant since a refocusing slice can be obtained directly from the micro image representation.

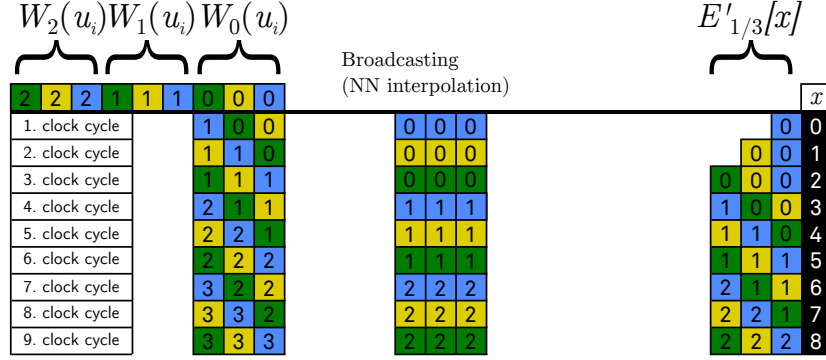


Figure 6. Horizontal shift and integration data flow diagram for sub-pixel shift  $a = 1/3$

Issues arise with a semi-systolic array when the number of filter taps increases as the output has to be broadcasted to more registers. For instance, when having large micro images, the filter length  $N$  becomes longer as it equals the micro image resolution  $m$ . The more registers required, the longer the wires get, in order to reach registers being far away from the respective output register. In electronics it is well known that long wires involve low-pass filter behaviour which leads to a roll-off factor on the given signal, making it difficult to keep the process synchronized. Usually, contrary to a semi-systolic arrangement, the systolic filter contains an additional delay register in each of its taps to prevent broadcasting on long wires. Therefore, it is useful to implement delays in the broadcasting net to avoid desynchronisation.

Although an FIR filter diagram for a shift by one sub-pixel is straightforward, higher sub-pixel shifts appear to be more sophisticated. This is due to the fact that only by doing a sub-pixel shift of  $a = 1/3$  neighbored pixel positions of different micro images are merged together. To paraphrase, the gap between pixels from different micro images being averaged varies when shifting  $> 1/m$ .

$$E'_{2/3}[x]_{(1)} = \frac{1}{3} \left( E_{fs}[x-0] + E_{fs}[x-1] + E_{fs}[x-2] \right) \quad (23)$$

$$E'_{2/3}[x]_{(2)} = \frac{1}{3} \left( E_{fs}[x-1] + E_{fs}[x-2] + E_{fs}[x-3] \right) \quad (24)$$

$$E'_{2/3}[x]_{(3)} = \frac{1}{3} \left( E_{fs}[x-2] + E_{fs}[x-3] + E_{fs}[x-4] \right) \quad (25)$$

There are three equations in (23), (24) and (25) representing three different stages for the synthesis of the horizontal refocusing slice  $E'_{2/3}[x]$ , since the shift and integration by a value  $a = 2/3$  involves averaging the illuminance of pixels which are not necessarily direct neighbors. To assemble these three equations in a FIR design and cope with the circumstance that pixels being in between a gap have to be omitted, adaptive switches are introduced to the inputs of the FIR adders. As shown in Figure 7, the sub-pixel shift  $a = 2/3$  requires three states of the switch setting, depending on the incoming data  $E[x]$ .



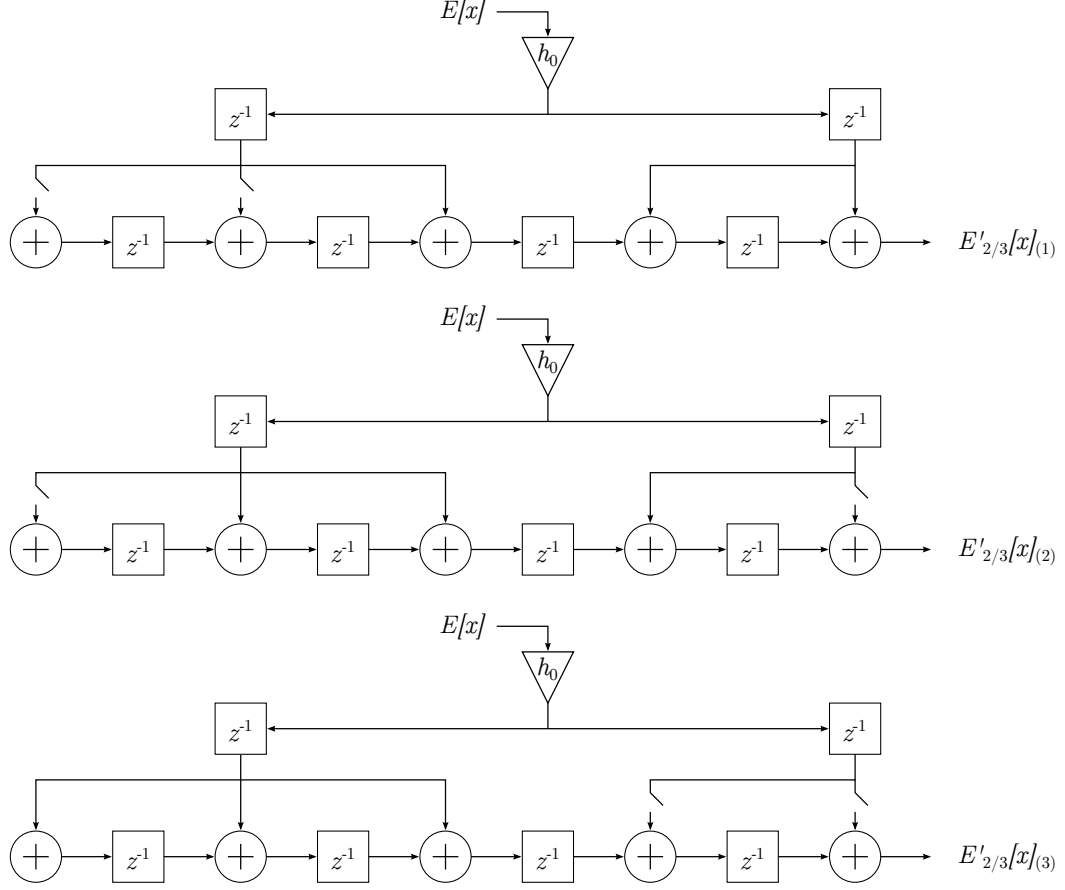


Figure 7. Three-state FIR filter diagram for horizontal sub-pixel shift  $a = 2/3$

When choosing a systolic arrangement, the length of the wires for broadcasting need to be taken into consideration. Due to the switch controlled design, applying the systolic arrangement that encompasses additional delays in each cell, would have a negative impact on the output pixels. To overcome the problem of long wires in the broadcasting net, a tree structure of intermediate registers is implemented, serving the incoming pixel value to a fixed number of adders in the semi-systolic processing part. Figure 8 illustrates the data flow diagram for a sub-pixel shift  $a = 2/3$  which again does not consider the fraction process. Note that pixels at the output  $E'_{2/3}[x]$  are delayed by one clock cycle because of the additional delay registers.

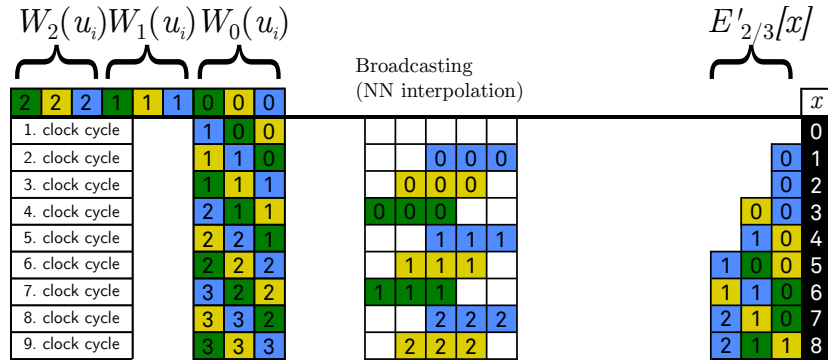


Figure 8. Horizontal shift and integration data flow diagram for sub-pixel shift  $a = 2/3$

### 3.1.2 Vertical shift and integration

After processing the shift and integration in horizontal direction, the pixel subsets  $E'[x, y]$  of parallel frame lines  $y$  are pipelined to the vertical shift and integration processor block. For a vertical sub-pixel shift  $a = 1/3$  the FIR filter design is exactly the same as in horizontal direction, assuming the micro images to be square sized. The only distinction is that horizontally processed pixels  $E'[x, y]$  with the same index  $x$  of several horizontal lines  $y$  have to be averaged. The vertical FIR filter equation

$$E''_{1/3}[x, y] = \frac{1}{3} \left( E'_{1/3}[x, y - 0] + E'_{1/3}[x, y - 1] + E'_{1/3}[x, y - 2] \right), \quad (26)$$

for sub-pixel shift  $a = 1/3$  indicates a similar arrangement.

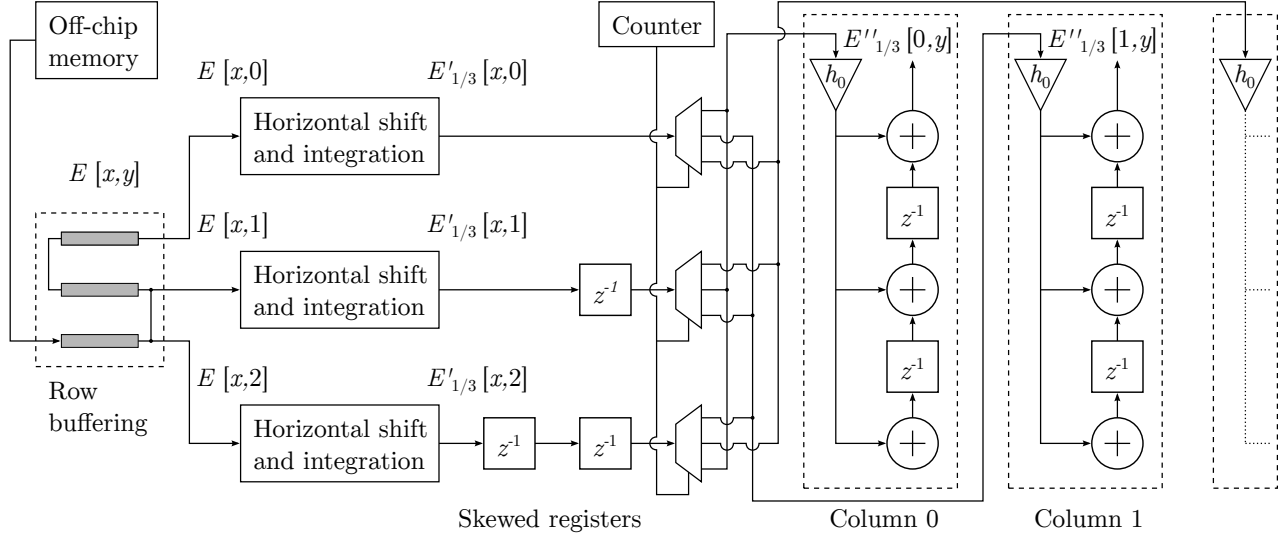


Figure 9. Parallelized semi-systolic FIR filter for vertical sub-pixel shift  $a = 1/3$

As depicted in Figure 9, to enable the parallelization, an array of demultiplexers has been implemented, assigning the pixels  $x$  of lines  $y$  to a vertical shift and integration FIR filter block correspondingly. The demultiplexers are driven by a counter. To prevent loading too many pixel values into a filter block at the same time, the parallelized horizontal lines are delayed in a skewed manner. The FIR blocks in column 0 and 1 are equivalent and the number of blocks required corresponds to the vertical resolution of the input image. When applying the parallelization scheme to other shift values, the FIR blocks can be simply replaced by FIR filters of higher shift values. The two-dimensional processed output  $E''[x, y]$  is buffered and send to off-chip memory afterwards.

## 3.2 Results

The original images provided by our camera have a resolution of  $5616 \times 3744$  pixels. The structure of the MLA is specified as rectangular, consisting of 191 by 126 micro lenses. As the clock frequency of the XC6SLX45 FPGA restricts the frame size to HD ready (720p) video standard, the images have been preprocessed by cropping each micro image appropriately. Thereby, after cropping the micro image size amounts to  $m = 6$  in horizontal and vertical dimension, vignetting disappears and the aligned input resolution amounts 1140 by 720 pixels matching the given hardware requirements of the FPGA.

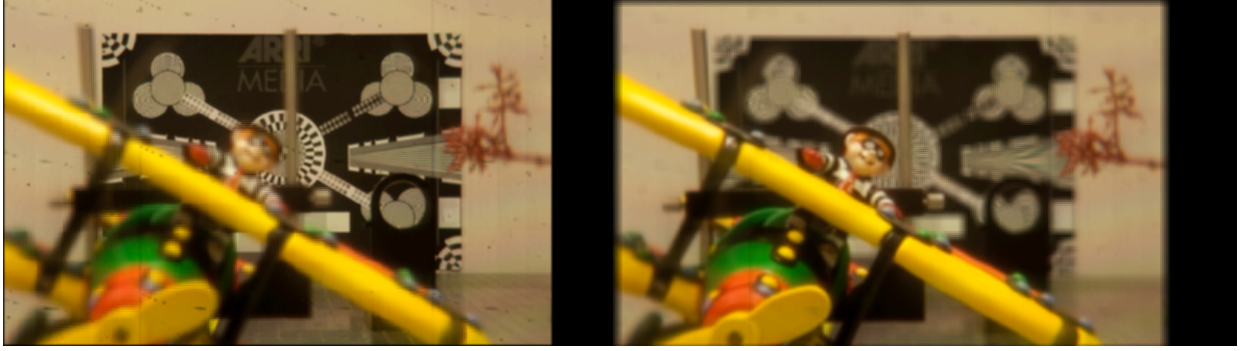


Figure 10. Refocusing results (left: sub-pixel shift  $a = 1/6$  , right: sub-pixel shift  $a = 4/6$  )

Examples of the refocused slices are demonstrated in Figure 10. As displayed in the images, the image synthesis provides refocusing slices at several distances by varying the parameter of the sub-pixel shift. Due to the simplicity of nearest-neighbor interpolation, the effective resolution of those images being sub-pixel shifted by a multiple of the micro image resolution  $m$ , more specifically shifted by an integer pixel, is effectively the same as it would be without interpolation. In that particular case, this is due to the fact that repeated pixel values of different samples  $x$  do not overlap each other in the shift and integration process. An additional stage for a better interpolation method is desirable for future work as this will smooth the images at critical sub-pixel shifts.

#### 4. CONCLUSION

In this investigation, the aim was to design a hardware implementation to synthesize a refocusing slice on the basis of a plenoptic camera. Returning to the hypothesis posed at the beginning, it is now possible to state that the refocusing slice synthesis does not necessarily include the sub-aperture extraction as an intermediate stage. Furthermore, the method used in this project has shown that a direct synthesis from micro images simplifies the process by using switch controlled FIR filters. To the best of our knowledge, this paper presents the very first hardware-based approach enabling the generation of two-dimensional focused slices of a light field, acquired by a plenoptic camera.

In the current setup, the Spartan 6 XC6SLX45 FPGA restricts the resolution to HD ready as its clock frequency is limited. Nevertheless, applying interpolation techniques, an output resolution of 720p real-time content is achievable with cameras capable of streaming according to the HDMI standard.

The generated images suffer from artifacts which are a result of refocusing a single slice that is parallel to the image sensor. These artifacts are accepted as the given design is meant for real-time preview purposes to assist cinematographers in their assessment. However, it might be an interesting starting point for a future research project to route the input content from the camera source to several processors of the proposed architecture to compute all possible refocusing slices in parallel. As previously elaborated by other researchers, it is feasible to combine these slices using a depth map to form an all-in-focus image with an extended depth of field.

#### REFERENCES

- [1] Ives, F. E., “Parallax stereogram and progress of making same,” (April 1903).
- [2] Lippmann, M. G., “Épreuves réversibles donnant la sensation du relief,” *Académie Des Sciences* , 446–451 (March 1908).
- [3] Sokolov, P. P., “Autostereoscopy and integral photography by prof. lippman’s method,” tech. rep. (1911).
- [4] Ives, H. E., “Parallax panoramagrams made with a large diameter lens,” *Journal of the Optical Society of America* **20**, 332–340 (June 1930).
- [5] Coffey, D. F., “Apparatus for making a composite stereograph,” Tech. Rep. 2063985, U.S. Patent Commission (December 1936).

- [6] Adelson, E. H. and Bergen, J. R., “The plenoptic function and the elements of early vision,” *Computational Models of Visual Processing*, 3–20, Cambridge, MIT Press (1991).
- [7] Levoy, M. and Hanrahan, P., “Light field rendering,” tech. rep., Stanford University (1996).
- [8] Isaksen, A., *Dynamically Reparameterized Light Fields*, Master’s thesis, Electrical Engineering and Computer Science, Massachusetts Institute of Technology (November 2000).
- [9] Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., and Hanrahan, P., “Light field photography with a hand-held plenoptic camera,” Tech. Rep. CTSR 2005-02, Stanford University (2005).
- [10] Wimalagunaratne, R., Madanayake, A., Dansereau, D., and Bruton, L., “A systolic-array architecture for first-order 4-d iir frequency-planar digital filters,” 3069–3072, ISCAS 2012 (May 2012).
- [11] Lumsdaine, A. and Georgiev, T., “Full resolution lightfield rendering,” tech. rep., Adobe Systems, Inc. (January 2008).
- [12] Ng, R., *Digital Light Field Photography*, PhD thesis, Stanford University (July 2006).
- [13] Perwaß, C. and Wietzke, L., “Single lens 3d-camera with extended depth-of-field,” in [*Human Vision and Electronic Imaging XVII*], **Proc. SPIE 8291**, Raytrix GmbH (February 2012).
- [14] Bishop, T. and Favaro, P., “Plenoptic depth estimation from multiple aliased views,” in [*3DIM 09*], (2009).
- [15] Bishop, T. and Favaro, P., [*Full-Resolution Depth Map Estimation from an Aliased Plenoptic Light Field*], pp 186–200, 6493, Springer Berlin Heidelberg (November 2010).
- [16] Steurer, J. H., Pesch, M., and Hahne, C., “3d holoscopic video imaging system,” in [*Human Vision and Electronic Imaging XVII*], **Proc. SPIE 8291**, Arnold & Richter GmbH & Co. KG (February 2012).
- [17] Aggoun, A. e. a., “Immersive 3d holoscopic video system,” *IEEE Multimedia Special Issue: 3D Imaging Techniques and MultiMedia Applications* **20(1)**, 28–37 (2013).
- [18] Rodríguez-Ramos, J., Lüke, J., López, R., Marichal-Hernández, J., Montilla, I., Trujillo-Sevilla, J., B. Femenía, M. P., López, M., Fernández-Valdivia, J., Rosa, F., Dominguez-Conde, C., Sanluis, J., and Rodríguez-Ramos, L., “3d imaging and wavefront sensing with a plenoptic objective,” in [*Three-Dimensional Imaging, Visualization, and Display*], **Proc. SPIE 8043**, Universidad de La Laguna (May 2011).
- [19] Dansereau, D. G., Pizarro, O., and Williams, S. B., “Decoding, calibration and rectification for lenselet-based plenoptic cameras,” in [*Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference*], 1027 – 1034 (June 2013).
- [20] Johannsen, O., Heinze, C., Goldluecke, B., and Perwaß, C., “On the calibration of focused plenoptic cameras,” in [*GCPR Workshop on Imaging New Modalities*], (September 2013).
- [21] LCC, *MHL-Enabled Devices*. <http://www.mhlconsortium.org/> (July 2013).
- [22] Silicon Image, Inc., *High-Definition Multimedia Interface*, version 1.2 ed. (August 2005).
- [23] Bailey, D. G., [*Design for Embedded Image Processing on FPGAs*], Wiley-Blackwell (2011).
- [24] Rodríguez-Ramos, L., Marín, Y., Díaz, J., Piqueras, J., García-Jiménez, J., and Rodríguez-Ramos, J., “Fpga-based real time processing of the plenoptic wavefront sensor,” Instituto de Astrofísica de Canarias, Santa Cruz de Tenerife, EDP Sciences (February 2010).
- [25] Fatah, O., e. a., “Three-dimensional integralimage reconstruction based on viewpoint interpolation,” in [*Proceedings of IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*], (2013).